LLNL-SR-654064

LAWRENCE LIVERMORE NATIONAL LABORATORY

Metal Artifact Reduction for CT-based Luggage Screening

S. Karimi, P. Cosman

May 8, 2014

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Metal Artifact Reduction for CT-based Luggage Screening

Seemeen Karimi and Pamela Cosman

Abstract

This report describes research on active problems in computed tomography (CT) scanning applications. The problems are metal artifact reduction in CT images for medical imaging and luggage screening, and methods to evaluate segmentation of CT images. The research forms the basis of a PhD dissertation topic at UCSD, and is sponsored by LLNL.

I. INTRODUCTION In aviation security, luggage screening is often done by CT-based screening systems, which employ automatic target recognition algorithms (ATR). The U.S. Department of Homeland Security has identified lowering false alarms and increasing threat categories as a requirement for future systems. This motivates improvements in image reconstruction, image segmentation and ATR [1]. The original goal of this research was the recognition and characterization of ordinary nonthreat objects in luggage, encompassing image reconstruction, segmentation and the evaluation of segmentation. The presence of metal in luggage creates artifacts which are a large impediment in segmentation. Therefore, research in metal artifact reduction (MAR) was established as an important goal. Most of the MAR literature comes from medical imaging. We made improvements in MAR for the medical application. The medical work helped us better understand the metal artifact problem and to develop a solution for luggage scanning. We also developed segmentation evaluation methods suited for luggage screening, and that give insights into the machine segmentation algorithms that were previously unavailable from other evaluation methods. Finally, we reviewed existing literature on segmentation algorithms, and implemented promising methods.

II. METAL ARTIFACT REDUCTION (MAR) IN MEDICAL CT The presence of metal in CT scans causes streaks and shadows that obscure surrounding tissue, making it difficult for radiologists to evaluate images. For over three decades MAR has been an active area of research. The various approaches fall into three categories, sinogram replacement [2]–[10], multiple-energy decomposition [11]–[17] and iterative reconstruction [12], [18]–[23]. We have worked on a sinogram replacement method because it is faster than numeric reconstruction, and because scanning with two or more energies is not standard scanning practice. In the sinogram replacement approach, projection samples in the sinogram (Radon space) corresponding to rays that pass through metal are replaced with an estimate of true underlying data. The rays are calculated from an original image reconstruction that contains artifacts. In recent years, it was proposed that an intermediate coarse image be reconstructed, which could then be reprojected and guide the replacement of the metal-contaminated samples [5], [6]. This image is often called a ”prior-image”. The prior image is generated by the voxel-wise classification of the original image into tissue types. The classification is done by thresholding. Thresholding often results in misclassification of tissue leading to residual artifacts or secondary artifacts. Our advance on this approach consists of making a better prior-image by exploiting observations we have made about the characteristics of metal artifacts; the artifacts are adjacent to metals, artifact intensity drops with distance from the metal, and that local maxima in Radon space correspond to local minima in the image space. These observations are based on simulations that we performed. We segment the artifacts from anatomy using morphology and clustering, using the properties given above. We replace the artifact-labeled areas with soft tissue values. The prior image is reprojected and used with [5] to obtain corrected data. The corrected data is reconstructed to give the final MAR image. Since spatial and CT intensity

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

distributions are jointly used, we can get a better segmentation than by using intensity (thresholding) alone.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

We applied our method to medical scans of heads obtained from Lahey Clinic, Burlington MA. Eight images were tested with metal implants including aneurysm coils, a deep brain stimulator and dental fillings. Our MAR images showed good artifact reduction. We implemented other methods in the literature for comparison [3], [24], and found that our method yielded better artifact reduction than the others, and the improvement came solely from the prior. We showed that the prior has more impact on the final image than the data replacement technique. Our work opens up an interesting area of research, which can be extended for better artifact reduction, robustness to more metal and for application to other anatomical areas.

This work led to peer-reviewed conference and journal publications. Details of our methods and results can be found in the journal paper.

S. Karimi, P. Cosman, C. Wald, and H. Martz, "Segmentation of artifacts and anatomy in CT metal artifact reduction," Medical Physics, vol. 39, pp. 585768, 2012.

S. Karimi, P. Cosman, C. Wald, and H. Martz, "Using segmentation in CT metal artifact reduction," In IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), 2012 , pp. 9-12. IEEE, 2012.

III. Metal Artifact Reduction in CT-based Luggage Screening

Similar to the medical application, metal artifacts degrade the CT images making it difficult for ATR algorithms to correctly segment and recognize objects of interest. The artifacts may lead to apparent splitting of a single object, or the merging of separate objects. Since EDS are tuned for high detection rates, the artifact-degraded images lead to higher false alarms. Reducing metal artifacts is expected to improve system performance. In luggage scanning, the contents of the bags are unknown. Therefore, the sinogram replacement techniques mentioned in the previous section cannot be used because they create prior-images through image segmentation, and segmentation is based on assumptions about image contents. Therefore, we look to iterative and numerical techniques. Model based iterative techniques have the potential to reduce metal artifacts but rely on the accuracy of attenuation process, which are difficult to model correctly and are slow. They also often require that the scan materials be known [20], [22], [25], [26]. A recent approach to MAR is to use numerical optimization for reconstruction without detailed scanner modeling. This approach assumes that the projection data are adequately preprocessed to compensate for other image degradations, but are still degraded by metal. Numerical optimization has become more reliable and efficient in recent years, but its application to MAR is limited [27]–[29]. These methods use different objective functions and constraints in their methods, but in all, the sinogram samples containing metal are discarded. As a result, metal artifacts are deleted, but there is a loss of edges. Our approach is again to build a prior-image, but without the assumptions from the medical application, and to use the prior-image in sinogram replacement. Like the medical application, we want our prior image to have sparse gradients and be artfact-free. We do not discard metal projection samples, but rather, we deemphasize them by using a weighting function. We perform a constrained regularized weighted least squares minimization. We use total variation regularization following earlier methods [28], [30]. We choose exponential weights because attenuation is exponential, and the weights are smooth and monotonic. We have added a novel constraint, which reduces the possibility of artifacts being pushed elsewhere due to the weighting function. Our constraint is that reprojected rays through metal must be greater than the rays measured from the scanner. A third innovation was to shrink the optimization problem. The artifacts that are difficult to correct are large low-frequency artifacts. By taking the difference between our constrained optimization solution, and an unweighted least-squares solution we isolated the artifacts. We solved for

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

images that were $\frac{1}{16}$ the original size and downsampled the projection data by another factor of 16, allowing a speed-up of 16^3 in reconstruction time. The isolated artifacts were upsampled and subtracted from the original image to yield the prior-image. The prior-image was then used with [5] to correct the data, and the corrected data reconstructed to give the final image. We tested our method on eight bag images obtained from a medical scanner (Imatron, CA), courtesy the ALERT group at Northeastern University. The bags were packed with various objects and different kinds of metallic objects, and contained various levels of clutter. In visual and quantiative evaluation, our method provided good artifact reduction. Objects with uniform CT attenuation, such as contained liquids were present in each bag. The CT number distributions within these objects were used for quantitative evaluation. We also used gradient-based scores, and sinogram-based errors defined in the literature for evaluation. A limitation is that edges were lost when the streaks were along them, which is common to most MAR methods. We implemented other methods that were applicable to non-medical scanning [3], [28] for comparison. Our method yielded better results than these benchmarks in that edges were retained better, and fewer secondary artifacts resulted. Details are provided in a paper that is submitted to a peer-reviewed journal. We have also submitted this research to a conference.

S. Karimi, H. Martz, and P. Cosman, "Metal Artifact Reduction for CT-based Luggage Screening," submitted to IEEE Transactions on Image Processing

S. Karimi, H. Martz, and P. Cosman, "Metal Artifact Reduction for CT-based Luggage Screening," submitted to IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2014.

IV. Segmentation Evaluation

Quantitative evaluation is necessary to assess machine segmentation (MS) algorithms meaningfully. For applications such as luggage scanning, CT images contain many objects for an ATR to segment and characterize. Therefore, the evaluation method should provide useful results with multiple segments, for multiple split and merge errors. We have several requirements of the evaluation algorithm. First, the evaluation method should provide insight into the behavior of the MS algorithm, so that the latter can be improved, e.g., tendency toward over / undersegmentation. Second, the method must evaluate the extent to which an MS algorithm captures object features. Third, it should be able to assign priorities to segments when evaluating the MS algorithm. Priorities may be based on image features. Various goodness measures (such as region similarity and inter-region differences) have been proposed to evaluate a segmentation without a ground truth (GT) reference [31]–[33]. However, luggage articles are inherently heterogeneous in composition, making goodness measures unsuitable in this application. Other methods evaluate segmentation against GT by computing a distance between the sets of edge pixels [31], [34], [35], but these methods do not measure feature retrieval. Local and global consistency errors from computer vision literature ignore refinements [36], which in our application correspond to label split and merge errors. Some other evaluation methods do not measure feature recovery [37], measure multiple features from single labels (in background) [38], or measure a total feature value from multiple labels [39], [40]. Other researchers have proposed measuring differences between histograms [41], [42] which makes sense when the objects of interest are similar and their features characterize populations. We propose two new methods of evaluation to meet the application needs described above, and address many limitations of existing methods. Both methods require GT segmentations. Our images are 3D images containing multiple objects that have complex shapes. To create GTs efficiently, we developed a semiautomatic labeling method by combining manual contouring, contour interpolation, and region growing. Both our evaluation methods require

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

that we generate a confusion matrix, whose rows consist of GT labels and columns consist of MS labels. The confusion matrix cells may contain the number of voxels common to the row-column pair, or contain the value of any pointwise feature. The first evaluation method is based on information theory. We calculate a weighted mutual information

(WMI) score of features from their joint distribution in GT and MS. The confusion matrix allows GT and MS label images to be expressed as joint and marginal probability densities. The mutual information is normalized by the square root of the product of the entropies. The confusion matrix can be weighted row-wise to emphasize certain objects or properties before computing the score. We have used the WMI score for volume, mass and mass weighted in a way to prioritize uniform objects. The mass and uniformity are computed with respect to the CT image.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

The second evaluation method, which we call Feature Descriptor Recovery (FDR), is based on establishing best correspondence between MS and GT segments and measuring segment-wise errors. The best correspondence was established using the Hungarian algorithm. As in the WMI score, the error can be computed for segment volumes, masses or uniformity (weighted by mass), or any other desired feature. We can determine whether the errors are predominantly undersegmentation, oversegmentation or random, as well as determine outliers and trends.

A database of CT images of suitcases was generated by the ALERT group at Northeastern University, and distributed to five research groups at universities and corporations [43]. The database contained no threats; the requirement was to segment all objects present in each suitcase. We obtained the results of the algorithms on five suitcases, and used our measures to evaluate the MS algorithms. Both evaluation methods have different perspectives, however their results were in agreement. Some additional findings were that all the MS algorithms did a better mass retrieval than volume retrieval, and that some algorithms may have trends, such as better accuracy for certain feature values. The evaluation methods were validated by human expert observer experiments on the bags, and by synthetic problems.

More details are available in the following journal and conference publications.

S. Karimi, X. Jiang, P. Cosman, H. Martz, "Flexible Methods for Segmentation Evaluation: Results from CT-based Luggage Screening," Journal of X-ray Science and Technology, accepted Jan 2014.

S. Karimi, X. Jiang, P. Cosman, H. Martz, "Evaluation of Segmentation Algorithms in CT scanning," In IEEE Second International Conference on Healthcare Informatics, Imaging and Systems Biology (HISB), 2012 , pp. 139-139. IEEE, 2012.

V. Segmentation Algorithm review

ATR algorithms segment objects of interest in luggage images and measure properties of those objects. Improved segmentation will lead to lower false alarm rates. In addition, the definition of threats is evolving, so ATRs can be improved by better characterization of all objects in bags. We have investigated and implemented several segmentation algorithms and weighed benefits and weaknesses for CT-based luggage screening. We have reviewed level set methods [44], [45], Markov random fields [46], graph-cut (GC) methods [47]–[51], random walks [52], mean shift [53], region growing [54], watershed segmentation [55], conditional random fields (CRF) [56] and hierarchical clustering methods [56]–[59]. We have implemented region growing, watershed segmentation and graph cut methods. We believe GC methods are the most promising for segmentation in luggage. CRF is another promising approach because it allows supervision in segmentation. Supervision, such as training for compatibility of pixels or superpixels is important for this application because the objects are diverse and heterogenous. We have implemented a segmentation algorithm based on GC. We used expectation maximization to determine the prior probabilites of the nodes, and the conditional probabilities, assuming a Gaussian mixture model (GMM). We tested this EM/GC segmentation algorithm on simulated images and a bag slice. The simulated data gave us a good segmentation, but the real data case combined some segments that belonged to different objects because they had similar characteristics.

Additional study is needed for better estimation of model order, optimization, building contexts for hierarchical segmentation.

References

[1] "Final Report on Algorithm Devlopment for Security Applications," in *Algorithm Development for Security Applications*, M. Silevitch, C. Crawford, and H. Martz, Eds., Boston, 2009.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

[2] G. Glover and N. Pelc, "An algorithm for the reduction of metal clip artifacts in CT reconstructions," *Medical Physics*, vol. 8, pp. 799–807, 1981.

[3] W. Kalender, R. Hebel, and J. Ebersberger, "Reduction of CT artifacts caused by metallic implants," *Radiology*, vol. 164, no. 2, pp. 576–77, 1987.

[4] S. Zhao, D. D. Robertson, G. Wang, B. Whiting, and K. T. Bae, "X-ray CT metal artifact reduction using wavelets: an application for imaging total hip prostheses." *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1238–47, Dec. 2000.

[5] R. Naidu, I. Bechwati, S. Karimi, S. Simanovsky, and C. Crawford, "Method of and system for reducing metal artifacts in images generated by x-ray scanning devices," U.S. Patent 6 721 387, 2004.

[6] M. Bal and L. Spies, "Metal artifact reduction in CT using tissue-class modeling and adaptive prefiltering," *Medical Physics*, vol. 33, no. 8, pp. 2852–2859, Jul. 2006.

[7] C. Golden, S. R. Mazin, F. E. Boas, G. Tye, P. Ghanouni, G. Gold, M. Sofilos, and N. J. Pelc, "A comparison of four algorithms for metal artifact reduction in CT imaging," in *SPIE Medical Imaging*, N. J. Pelc, E. Samei, and R. M. Nishikawa, Eds. International Society for Optics and Photonics, Mar. 2011, p. 79612Y.

[8] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, "Frequency split metal artifact reduction (FSMAR) in computed tomography." *Medical Physics*, vol. 39, no. 4, pp. 1904–16, Apr. 2012.

[9] F. E. Boas and D. Fleischmann, "Evaluation of two iterative techniques for reducing metal artifacts in computed tomography." *Radiology*, vol. 259, no. 3, pp. 894–902, Jun. 2011.

[10] T. Koehler, B. Brendel, and K. Brown, "A New Method for Metal Artifact Reduction in CT," in *The International Conference in X-ray Computed Tomography*, Salt Lake City, Utah, USA, 2011.

[11] R. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," *Physics in Medicine and Biology*, vol. 21, no. 5, pp. 733–744, 1976.

[12] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "An iterative maximum-likelihood polychromatic algorithm for CT." *IEEE Transactions on Medical Imaging*, vol. 20, no. 10, pp. 999–1008, Oct. 2001.

[13] F. Bamberg, A. Dierks, and K. Nikolaou, "Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation," *European Radiology*, vol. 21, no. 7, pp. 1424–1429, 2011.

[14] P. Sukovic and N. Clinthorne, "Penalized weighted least-squares image reconstruction for dual energy X-ray transmission tomography," *IEEE Transactions on Medical Imaging*, vol. 19, no. 11, pp. 1075 – 1081, 2000.

[15] H. Xue, L. Zhang, and Y. Xiao, "Metal artifact reduction in dual energy CT by sinogram segmentation based on active contour model and TV inpainting," in *Nuclear Science Symposium Conference Record (NSS/MIC)*, Orlando, FL, 2009, pp. 904–908.

[16] H. Li, L. Yu, X. Liu, J. G. Fletcher, and C. H. McCollough, "Metal artifact suppression from reformatted projections in multislice helical CT using dual-front active contours," *Medical Physics*, vol. 37, no. 10, p. 5155, Sep. 2010.

[17] C. Zhou, Y. Zhao, S. Luo, H. Shi, and L. Zheng, "Monoenergetic imaging of dual-energy CT reduces artifacts from implanted metal orthopedic devices in patients with factures," *Academic Radiology*, vol. 18, no. 10, pp. 1252–1257, 2011.

[18] G. Wang and D. Snyder, "Iterative deblurring for CT metal artifact reduction," *IEEE Transactions on Medical Imaging*, vol. 15, no. 5, pp. 657–664, 1996.

[19] M. Oehler and T. Buzug, "Modified MLEM algorithm for artifact suppression in CT," in *Nuclear Science Symposium Conference Record*, vol. 6, San Diego, CA, 2006, pp. 3511–3518.

[20] R. Murphy, S. Yan, J. O'Sullivan, D. Snyder, B. Whiting, D. Politte, G. Lasio, and J. Williamson, "Pose estimation of known objects during transmission tomographic image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 25, no. 10, pp. 1392–1404, Oct. 2006.

[21] C. Lemmens, D. Faul, and J. Nuyts, "Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion," *IEEE Transactions on Medical Imaging*, vol. 28, no. 2, pp. 250–260, 2009.

[22] J. Stayman and Y. Otake, "Model-Based Tomographic Reconstruction of Objects Containing Known Components," *IEEE Transactions on Medical Imaging*, vol. 31, no. 10, pp. 1837–1848, 2012.

[23] T. Sakimoto and K. Nishino, "Metal artifact reduction in tomosynthesis by metal extraction and ordered subset-expectation maximization (OS-EM) reconstruction," in *SPIE Medical Imaging*, 2013, p. 86685M.

[24] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelriess, "Normalized metal artifact reduction (NMAR) in computed tomography." *Medical Physics*, vol. 37, no. 10, pp. 5482–93, Oct. 2010.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

[25] I. A. Elbakri and J. A. Fessler, “Statistical image reconstruction for polyenergetic X-ray computed tomography.” *IEEE Transactions on Medical Imaging*, vol. 21, no. 2, pp. 89–99, Feb. 2002.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

[26] ——, “Segmentation-free statistical image reconstruction for polyenergetic x-ray computed tomography with experimental validation,” *Physics in Medicine and Biology*, vol. 48, no. 15, pp. 2453–2477, Aug. 2003.

[27] X. Zhang, J. Wang, and L. Xing, “Metal artifact reduction in x-ray computed tomography (CT) by constrained optimization,” *Medical Physics*, vol. 38, no. 2, p. 701, 2011.

[28] J. M. Verburg and J. Seco, “CT metal artifact reduction method correcting for beam hardening and missing projections.” *Physics in Medicine and Biology*, vol. 57, no. 9, pp. 2803–18, May 2012.

[29] Y. Zhang, H. Yan, X. Jia, J. Yang, S. Jiang, and X. Mou, “A hybrid metal artifact reduction algorithm for x-ray CT,” *Medical Physics*, vol. 40, p. 041910, 2013.

[30] E. Y. Sidky and X. Pan, “Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization.” *Physics in Medicine and Biology*, vol. 53, no. 17, pp. 4777–807, Sep. 2008.

[31] Y. J. Zhang, “A survey on evaluation methods for image segmentation,” *Pattern Recognition*, vol. 29, no. 8, pp. 1335–1346, 1996.

[32] H. Zhang, J. E. Fritts, and S. A. Goldman, “Image segmentation evaluation: A survey of unsupervised methods,” *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 260–280, May 2008. [Online]. Available: http://dx.doi.org/10.1016/j.cviu.2007.08.003

[33] B. Johnson and Z. Xie, “Unsupervised image segmentation evaluation and refinement using a multi-scale approach,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 4, pp. 473–483, Jul. 2011. [Online]. Available: http://dx.doi.org/10.1016/j.isprsjprs.2011.02.006

[34] A. Fenster and B. Chiu, “Evaluation of Segmentation algorithms for Medical Imaging.” *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, vol. 7, pp. 7186–9, Jan. 2005. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=1616166

[35] T. Heimann, B. van Ginneken, M. a. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes, F. Bello,

G. Binnig, H. Bischof, A. Bornik, P. M. M. Cashman, Y. Chi, A. Cordova, B. M. Dawant, M. Fidrich, J. D. Furst, D. Furukawa,

L. Grenacher, J. Hornegger, D. Kainm¨uller, R. I. Kitney, H. Kobatake, H. Lamecker, T. Lange, J. Lee, B. Lennon, R. Li, S. Li, H.-P. Meinzer, G. Nemeth, D. S. Raicu, A.-M. Rau, E. M. van Rikxoort, M. Rousson, L. Rusko, K. a. Saddi, G. Schmidt, D. Seghers,

A. Shimizu, P. Slagmolen, E. Sorantin, G. Soza, R. Susomboon, J. M. Waite, A. Wimmer, and I. Wolf, “Comparison and evaluation of methods for liver segmentation from CT datasets.” *IEEE transactions on medical imaging*, vol. 28, no. 8, pp. 1251–65, Aug. 2009. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19211338

[36] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 2. IEEE Comput. Soc, 2001, pp. 416–423. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=937655

[37] J. S. Cardoso and L. Corte-Real, “Toward a generic evaluation of image segmentation.” *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 14, no. 11, pp. 1773–82, Nov. 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16279178

[38] Y. Zhang and J. Gerbrands, “Segmentation evaluation using ultimate measurement accuracy,” *Proceedings of SPIE*, 1992. [Online]. Available: http://scholar.google.com/scholar?cluster=12715242476050196199&hl=en&as sdt=1,5#0

[39] F. Monteiro and A. Campilho, “Distance measures for image segmentation evaluation,” *AIP Conference Proceedings*, 2012. [Online]. Available: http://link.aip.org/link/?APCPCS/1479/794/1

[40] V. Mezaris, I. Kompatsiaris, and M. Strintzis, “Still Image Objective Segmentation Evaluation Using Ground Truth,” in *5th COST 276 Workshop*, B. Kovar, J. Prikryl, and M. Vlcek, Eds., 2003, pp. 9–14.

[41] C. Hagwood and J. Bernal, “Evaluation of Segmentation Algorithms on Cell Populations Using CDF Curves,” *Medical Imaging, IEEE ...* , 2011. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=6029989

[42] Y. Rubner, J. Puzicha, C. Tomasi, and J. M. Buhmann, “Empirical Evaluation of Dissimilarity Measures for Color and Texture,” *Computer Vision and Image Understanding*, vol. 84, no. 1, pp. 25–43, Oct. 2001. [Online]. Available:

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

http://dx.doi.org/10.1006/cviu.2001.0934

[43] "Final report on Algorithm Development for Security Applications 6," in *Algorithm Development for Security Applications*, M. Silevitch,

C. Crawford, and H. Martz, Eds. Boston: Northeastern University, 2011.

[44] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations," *Journal of computational physics*, vol. 79, no. 1, pp. 12–49, 1988.

[45] T. F. Chan and L. A. Vese, "Active contours without edges ," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

[46] S. Geman and D. Geman, “ Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images,” *Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–41, 1984.

[47] Z. Wu and R. Leahy, “An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 11, pp. 1101–1113, 1993.

[48] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, “A comparative study of energy minimization methods for markov random fields,” in *Computer Vision–ECCV 2006*. Springer, 2006, pp. 16–29.

[49] A. Delong, A. Osokin, H. N. Isack, and Y. Boykov, “Fast approximate energy minimization with label costs,” *International Journal of Computer Vision*, vol. 96, no. 1, pp. 1–27, 2012.

[50] S. Gould, J. Rodgers, D. Cohen, G. Elidan, and D. Koller, “Multi-class segmentation with relative location prior,” *International Journal of Computer Vision*, vol. 80, no. 3, pp. 300–316, 2008.

[51] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 888–905, 2000.

[52] L. Grady, “Random walks for image segmentation,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 11, pp. 1768–1783, 2006.

[53] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603–619, 2002.

[54] R. Gonzalez, R. Woods, and S. Eddins, *Digital image processing using MATLAB*, 1st ed. Pearson, 2009.

[55] L. Vincent and P. Soille, “Watersheds in digital spaces: an efficient algorithm based on immersion simulations,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 13, no. 6, pp. 583–598, 1991.

[56] J. Shotton, J. Winn, C. Rother, and A. Criminisi, “Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context,” *International Journal of Computer Vision*, vol. 81, no. 1, pp. 2–23, 2009.

[57] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “Contour detection and hierarchical image segmentation,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 5, pp. 898–916, 2011.

[58] S. C. Zhu and D. Mumford, *A stochastic grammar of images*. Now Publishers Inc, 2007, vol. 2, no. 4.

[59] X. He, R. S. Zemel, and D. Ray, “Learning and incorporating top-down cues in image segmentation,” in *Computer Vision–ECCV 2006*. Springer, 2006, pp. 338–351.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344